Housing Price Prediction Using Machine Learning

Prof. N. R. Jain[1], Prachi Zende[2], Mohini Gaikwad[3], Vibha Bhor[4]
[1]*HOD, Dept. Information Technology Engineering, P.D.E.A's college of Engineering, Pune, Maharashtra, India*
[2,3,4] *Student, Dept. Information Technology, Engineering, P.D.E.A's college of Engineering ,Pune, Maharashtra, India*

***Abstract*- House price forecasting is an important topic of real estate. The literature attempts to drive useful knowledge from historical data of property markets. Machine learning techniques are applied to analyze historical property transaction to discover useful models for house buyers and sellers. Revealed is the high discrepancy between house prices in the most expensive and most affordable suburbs. Moreover, experiments demonstrate that the combination of stepwise and support vector machine that is based on mean squared error measurement is a competitive approach.**

The goal of the study is through analyzing a real historical transactional dataset to derive valuable insight into the housing market. It seeks useful models to predict the value of a house given a set of its characteristics. Effective model could allow home buyers or real estate agents to make better decisions.

***Index terms*- Python, Spyder, Data integrity, Prediction, Data Modeling**

I.INTRODUCTION

Development of civilization is the foundation of increase of demand of houses day by day. Accurate prediction of house prices has been always a fascination for the buyers, sellers and for the bankers also. Many researchers have already worked to unravel the mysteries of the prediction of the house prices. There are many theories that have been given birth as a consequence of the research work contributed by the various researchers all over the world. Some of these theories believe that the geographical location and culture of a particular area determine how the home prices will increase or decrease whereas there are other schools of thought who emphasize the socio-economic conditions that largely play behind these house price rises. We all know that house price is a number from some defined assortment, so obviously prediction of prices of houses is a regression task. To forecast house price one person usually tries to locate similar properties at his or her neighbourhood and based on collected data that person will try to predict the house price. All these indicate that house price prediction is an emerging research area of regression which requires the knowledge of machine learning. This has motivated to work in this domain.

II. LITERATURE SURVEY

A) Prediction using various Machine Learning algorithms

There are two major challenges that researchers have to face. The biggest challenge is to identify the optimum number of features that will help to accurately predict the direction of the house prices. Kahn mentions that productivity growth in various residential construction sectors does impact the growth of the housing prices. The model that Kahn worked with shows how housing prices can have an apparently trendy appearance in which housing wealth rises faster than income for an extended period, then collapses and experiences an extended decline. Lowrance mentions in his doctoral thesis that he found the interior living space to be the most influential factor determining the housing prices with his research work. He also cites the medium income of the census tract that holds the prices. Pardoe utilizes features such as floor size, lot size category, number of bathrooms, and number of bedrooms, standardized age and garage size as features and utilizes linear regression techniques for predicting the house prices.

B) Forecasting the US Real House Price Index:

Structural and Non-Structural Models with and without Fundamentals Here employ10-variable

dynamic way structural general equilibrium model for forecast the US real house price index as well as its turning change point in 2006:Q2. Also there are various examines Bayesian and classical time-series models in our forecasting exercise for comparing to the dynamic stochastic general equilibrium model, by using Bayesian methods. In also to the addition on standard vector-autoregressive and Bayesian vector autoregressive models, and we also include the information content of either 10 or 120 quarterly series in few models for capturing the influence of fundamentals. Comparing the out-of-sample forecast performance of the alternative models, with use of the average root mean squared error for the forecasts. We conclude that the small-scale Bayesian-shrinkage model (10 variables) outperforms the other models, including the large-scale Bayesian-shrinkage model (120 variables). Finally at the end, we use each model to forecast the turning point in 2006:Q2, by using the estimated model through 2005:Q2.

C) Machine learning for a London housing price prediction mobile application.

Different regression methods were in the model which are explored and are compared. Gaussian Processes (GP) for regression was selected as our model because of its flexible and probabilistic approach towards learning and model selection criteria. In London for handling multiple types of the large dataset of the past property, the exploited structure of the dataset for distributing computations for the small independent local models. Overall obtained predictions are by recombining predictions from the current local models. And by training this models and obtaining predictions on the server-side of the application, then we are able to offload computationally intensive tasks performance and focus on generating visualisations on the client-side. Our results demonstrate that the approach of our towards the problem has been largely successful, and which is able to produce different predictions that are competitive to other housing price prediction models.

D) Housing Value Forecasting Based on Machine Learning Models

According to the various obtained predictions , the government and developers can make decisions about either developing the real estate market on corresponding regions or not.In this paper,support vector machines(SVM),least square support vector machines (LSSVM),and partial least square(PLS) methods are used for forecasting the house values. And these algorithms are compared on the basis of the predicted results. Previous history and experiments shows that although the data set have exists serious non-linearity, the experiment result also show SVM and LSSVM methods are superior to PLS on the dealing with problem of non-linearity. The global optimal solution can be found and best towards the forecasting effect can be achieved by SVM (support vector machine) because of solving a quadratic programming problem. In this paper, the multiple computation efficiencies of the algorithms are compared by according to the computing times related to relevant algorithms.

III.CONCLUSION

In the present real estate world, it has turned out to be difficult to store huge amount of information and concentrate them for one's own prerequisite. Likewise, the separated information ought to be helpful. The framework makes ideal utilization of all the models. It makes use of such information in the most effective way. The direct relapse calculation satisfies customer by expanding the exactness of their decision and diminishing the danger of putting resources into a home. A ton of highlights that could be added to make the framework all the more generally satisfactory.

REFERENCES

[1] Gupta, R., Kabundi, A., & Miller, S. M. (2011). Forecasting the US real house price index: Structural and non-structural models with and without fundamentals. Economic Modelling, 28(4), 2013-2021.

[2] Mu, J., Wu, F. & Zhang, A., 2014. Housing Value Forecasting Based on Machine Learning Methods. Abstract and Applied Analysis, 2014(2014), p.7.

[3] Ng, A., & Deisenroth, M. (2015). Machine learning for a London housing price prediction mobile application. Technical Report, June 2015, Imperial College, London, UK.

[4] Risse M. & Kern M., 2016. Forecasting house-price growth in the Euro area with dynamic

model averaging. North American Journal of Economics and Finance, 38, pp.70–85.

[5] Pino A 2018. Melbourne Housing Market data.Kaggle.https://www.kaggle.com/anthonypino/melbourne-housing-market.